DISCRETE MATHEMATICS 9 (1974) 109–117. © North-Holland Publishing Company

FIRST FACETS OF THE OCTAHEDRON

Fred GLOVER

Graduate School of Business, University of Colorado, Boulder, Colo. 80302, USA

Received 17 May 1973*

Abstract. A procedure is given for identifying the facets of the octahedron that are first intersected upon extending the edge of a polyhedral cone. The information generated by this procedure can be exploited to advantage by cut-search procedures for zero-one integer programming. Results are given which make it possible to determine the first two facets (or sets of "tied facets") following the innermost facet with less effort than required to determine the innermost facet itself. Depending on the orientation of the extended edge relative to the octahedron, a number of successive additional facets may be determined with comparable ease.

1. Introduction

The convexity (or intersection) cut defined relative to the octahedron requires the identification of the innermost facet encountered by an extended edge of a polyhedral cone (defined by a linear programming basis) whose vertex lies in the interior of the octahedron.

Such a procedure was the focus of the investigation of Balas, Bowman, Glover and Sommer [4].

However, it is desirable to have a way not only to determine the innermost facet encountered by an edge, but also to determine several subsequent facets, in the order in which they are encountered. The information thus generated can be exploited to advantage by cut-search procedures [5, 6].

This paper is devoted to identifying an efficient way to determine these first facets. Our approach requires less effort to determine the first two facets (or sets of "tied facets") following the innermost facet than required to determine the innermost facet itself (once its determina-

* Original version received 7 December 1971.

tion has been made). Depending on the orientation of the extended edge relative to the octahedron, a number of successive additional facets may be determined with comparable ease. A method is also provided that characteristically leads to a faster determination of the innermost facet than the method of [4].

Recently, Egon Balas [3] has devised a method for finding and ranking *all* facets of the octahedron based on a fundamental (and nonintuitive) result which says that each successive facet, after the first, must be adjacent to at least one facet previously intersected. (Balas also gives an improved procedure for determining the innermost facet.) The approach developed here complements that of Balas by showing that the problem of ranking the facets can be viewed in a different framework, and by giving a method specifically designed for dermining the first facets for applications, such as those of [6], in which knowledge of progressively deeper facets may be of less value than knowledge of facets that are closer to the interior surface. Our results are presented in the context of a "deformable octahedron" that is slightly more general than the octahedron discussed in [2, 3, 4].

2. Notation and definitions

Using translations (and optionally, integer combinations) of variables, the constraints

$$x = A_0 - At\,,$$
$$x \geq 0, \quad t \geq 0, \quad \text{and } x_i \text{ integer}, \quad i \in I,$$

of the mixed integer programming (MIP) problem give rise to a related set of constraints which we shall represent by

(1) $\qquad y = B_0 - Bt, \qquad t \geq 0\,,$

(2) $\qquad y_i \geq 1 \quad \text{or} \quad y_i \leq 0, \quad i \in P = \{1, ..., p\}.$

Consider any set of positive numbers h_i, $i \in P$, and let α denote a row vector of parameters each of whose components α_i, $i \in P$, is restricted to be 1 or -1. We associate with α the vector β given by

$$\beta_i = \begin{cases} 1 & \text{if } \alpha_i = 1, \\ 0 & \text{if } \alpha_i = -1, \end{cases}$$

and define the octahedron $K^*(h)$ to be the intersection of the 2^P half spaces

$$(3) \qquad H(\alpha, y) \equiv \sum_{i \in P} \alpha_i h_i y_i - \sum_{i \in P} \beta_i h_i \leq 0.$$

Thus, the hyperplanes $H(\alpha, y) = 0$ define the (extended) facets of $K^*(h)$. For a different characterization of the octahedron in terms of the L_1-norm (for the case in which all h_i are 1), see [3, 4].

Remark 2.1. All y contained in the unit hypercube $K = \{y\colon 1 \geq y_i \geq 0, i \in P\}$ must be contained in $K^*(h)$ and moreover, all y in K except for its $0-1$ vertices are in the interior of $K^*(h)$. In particular, given any α,

$$H(\alpha, y) < 0 \quad \text{for all } y \in K^*(h) \text{ except } y = \beta.$$

Under the assumption that B_0 is in the interior of $K^*(h)$ (as it will be if it is in K and not a vertex), our first goal, following [4], is to identify a value $\Delta^* > 0$ of t_j, such that the jth edge ($y = B_0 - B_j t_j$, $t_j \geq 0$) of the cone (1) intersects the boundary of $K^*(h)$ at $y^* = B_0 - B_j \Delta^*$. That is, defining $y(\Delta) = B_0 - B_j \Delta$, we seek the least positive value Δ^* of Δ such that $H(\alpha, y(\Delta)) = 0$ for some α (satisfying $\alpha_i = 1$ or -1). Viewing this problem in a slightly different manner, Δ^* is the least positive Δ such that $M(y(\Delta)) = 0$, where we define

$$M(y) = \max_{\alpha} \{H(\alpha, y)\}.$$

We shall call $M(y)$ the *critical displacement* of the point y from the boundary of the octahedron $K^*(h)$.

3. Some relations between hyperplanes defining $K^*(h)$ and critical displacements

The following results are a direct consequence of the definitions, and their proofs are omitted.

Lemma 3.1. *$M(y) = H(\alpha, y)$ if and only if*

$$\alpha_i = \begin{cases} 1 & \text{if } y_i > \frac{1}{2}\,, \\ -1 & \text{if } y_i < \frac{1}{2}\,, \\ 1 \text{ or } -1 & \text{if } y_i = \frac{1}{2}\,. \end{cases}$$

Lemma 3.2.

$$H(\alpha', y) - H(\alpha'', y) = \sum_{i \in P'} \alpha_i'' h_i(1-2y_i)\,,$$

where

$$P' = \{i \in P\colon \alpha_i' = -\alpha_i''\}.$$

Lemma 3.3. *$H(\alpha, y(\Delta'')) - H(\alpha, y(\Delta')) = (\Delta'' - \Delta')\,\theta(\alpha)$, where*

$$\theta(\alpha) = -\sum_{i \in P} \alpha_i h_i b_{ij}\,.$$

Lemma 3.4. *Let α' be given so that $H(\alpha', y(\Delta')) = M(y(\Delta'))$. Then if Δ'' is the solution to $H(\alpha', y(\Delta)) = 0$, $\Delta'' = \Delta' - M(y(\Delta'))/\theta(\alpha')$.*

Lemma 3.5. *Let α' be given so that $H(\alpha', y(\Delta)) = M(y(\Delta))$ for $\Delta = \Delta'$ and $\Delta = \Delta'' > \Delta'$. Then $M(y(\Delta'')) > M(y(\Delta'))$ if and only if $\theta(\alpha') < 0$.*

4. The innermost facet of $K^*(h)$

To determine the value Δ^* which gives the point $y(\Delta^*)$ at which the edge $y(\Delta)$ (for $\Delta \geq 0$) intersects the innermost facet of the octahedron, it follows immediately from Lemma 3.1 that we need only take into account the *breakpoint values* $(b_{i0} - \frac{1}{2})/b_{ij}$ ($i \in P$) of Δ which will cause $y_i(\Delta)$ (the ith component of $y(\Delta)$) to attain the value $\frac{1}{2}$. Only the breakpoint values that are positive and distinct are relevant, and we represent these values by

$$\gamma_1 < \gamma_2 < \ldots < \gamma_u, \quad u \leq p\,.$$

By convention, we also let $\gamma_0 = 0$ and $\gamma_{u+1} = \infty$. (If there are no positive breakpoint values, $u = 0$.) Then, by Lemma 3.1, there is an immediately identifiable $\alpha = \alpha'$ such that

$$M(y(\Delta)) = H(\alpha', y(\Delta))$$

for all Δ satisfying $\gamma_k \leq \Delta \leq \gamma_{k+1}$ $(0 \leq k \leq u)$. In particular, if y' is defined by $y' = y(\Delta)$ for any Δ such that $\gamma_k \leq \Delta < \gamma_{k+1}$, then α' may be given by

$$(4) \qquad \alpha_i' = \begin{cases} 1 & \text{if } y_i' > \tfrac{1}{2} \text{ or } y_i' = \tfrac{1}{2} \text{ and } b_{ij} \leq 0\,, \\ -1 & \text{otherwise.} \end{cases}$$

Equivalently, if y' is defined by $y' = y(\Delta)$ for any Δ satisfying $\gamma_k < \Delta \leq \gamma_{k+1}$, then α' can be given by

$$(5) \qquad \alpha_i' = \begin{cases} 1 & \text{if } y_i' > \tfrac{1}{2} \text{ or } y_i' = \tfrac{1}{2} \text{ and } b_{ij} \geq 0, \\ -1 & \text{otherwise.} \end{cases}$$

Note that (4) and (5) not only provide the same α'(for the indicated restrictions on $y' = y(\Delta)$), but these two definitions are indistinguishable for $\gamma_k < \Delta < \gamma_{k+1}$, in which case they result in setting $\alpha_i' = 1$ if $y_i' \geq \frac{1}{2}$ and $\alpha_i = -1$ if $y_i' < \frac{1}{2}$. Moreover, by Lemmas 3.3 and 3.5,

$$M(y(\gamma_{k+1})) = M(y(\gamma_k)) + (\gamma_{k+1} - \gamma_k)\, \theta(\alpha^k)\,,$$

where

$$\theta(\alpha^0) < \theta(\alpha^1) < \ldots < \theta(\alpha^u) = \theta(\alpha^{u+1}) = \sum_{i \in P} h_i\, |b_{ij}|$$

and α^k corresponds to the vector α' of (4) for $y' = y(\gamma_k)$. Thus, assuming $b_{ij} \neq 0$ for some i, there is a first positive γ_k, call it γ_r, such that $M(y(\gamma_r)) \geq 0$. Then, by the foregoing remarks, it follows that the desired Δ^* for which $M(y(\Delta^*)) = 0$ is the solution to the equation $H(\alpha', y(\Delta)) = 0$, where α' is determined from either (4) or (5) for $y' = y(\gamma_{r-1})$ or $y' = y(\gamma_r)$, respectively.

The value γ_r, and hence Δ^*, can be determined by successively computing and checking the γ_k values, as proposed in [4]. We note in this connection that the value of Δ^* can be computed with slightly less effort by reference to Lemma 3.4 than be directly solving for $H(\alpha', y(\Delta)) = 0$, the latter being essentially the proposal of [4]. Furthermore, the preceding observations provide another way to determine Δ^* without computing and ordering the breakpoint values, giving a more efficient approach for effecting this determination.

In particular, for *any* trial Δ' of Δ, one may identify a vector α' by Lemma 3.1 which specifies a *critical hyperplane* $H(\alpha', y')$ corresponding to the point $y' = y(\Delta')$; that is, a hyperplane for which the critical displacement of y' from the boundary of $K^*(h)$ is actually attained ($H(\alpha', y') = M(y')$). If this critical displacement $M(y')$ is negative, then α' may be given by (4), and if this displacement is positive, then α' may be given by (5). As a consequence, the solution Δ'' to $H(\alpha', y(\Delta)) = 0$ gives a new point $y(\Delta'')$ which is the intersection of the edge $y(\Delta)$ with the critical hyperplane. (The value Δ'' can be conveniently obtained in this approach by reference to Lemma 3.4. This value provides an upper bound for Δ^* whenever $\Delta'' > 0$, as will necessarily be the case if $M(y(\Delta')) > 0$.)

Thus, a simple and efficient procedure for determining Δ^* is to select any positive starting value for Δ', compute Δ'', and then replace Δ' with a new $\Delta' \leq \Delta''$ (unless $\Delta'' \leq 0$, in which case Δ' is simply increased).

Some analytical indication of the strength of this procedure can also be assessed. Specifically, if Δ' is an overestimate of Δ^* and $\gamma_k < \Delta' < \gamma_{k+1}$, $k \leq u$, then $\Delta'' \leq \gamma_k$ or $\Delta'' = \Delta^*$ (or both). This means that if Δ' is first selected sufficiently large, and successive values of Δ' are selected to equal Δ'', then usually fewer (and never more) iterations will result than by looking through the breakpoint values themselves (from γ_u down). (Also, if $\Delta^* \geq \gamma_u$, this procedure gives $\Delta'' = \Delta^*$ on the first step.)

5. First facets beyond the innermost facet

In order to identify successive facets encountered by the edge $y(\Delta)$ as Δ increases, we seek the successive values of $\Delta \geq \Delta^*$ such that $H(\alpha, y(\Delta)) = 0$ for some α. For each such value of Δ identified, we must determine the set of all α that yield $H(\alpha, y(\Delta)) = 0$, thereby identifying all facet hyperplanes which intersect the extended edge at a given point.

Theorem 5.1. *Let Δ' be an arbitrary (finite) positive value of Δ, and let γ_q be the smallest breakpoint value such that $\Delta' < \gamma_q$. Then the set of all facet hyperplanes $H(\alpha, y) = 0$ intersected by the edge $y = y(\Delta)$ for Δ in the interval $\Delta' \leq \Delta \leq \gamma_q$ is given by the set of all $0-1$ solutions to the equation*

$$\sum_{i \in P} d_i(\delta) z_i = M(y') + \delta\, \theta(\alpha')\,, \tag{6}$$

where $\theta(\alpha)$ is defined as in Lemma 3.3, α' is defined from (4) for $y' = y(\Delta')$, $\delta = \Delta - \Delta'$ and the coefficient $d_i(\delta)$ is given by

$$d_i(\delta) = \begin{cases} h_i(1-2(y_i'-\delta b_{ij})) & \text{if } i \in P_1 = \{i\colon y_i' < \frac{1}{2} \text{ or } y_i' = \frac{1}{2} \text{ and } b_{ij} \geq 0\}\,, \\ h_i(2(y_i'-\delta b_{ij})-1) & \text{if } i \in P_2 = P - P_1\,. \end{cases}$$

The $0-1$ solutions to (6) identify the facet hyperplanes $H(\alpha, y(\Delta)) = 0$ by the relationship

$$z_i = \begin{cases} 1 & \text{if } \alpha_i = -\alpha_i'\,, \\ 0 & \text{if } \alpha_i = \alpha_i'\,. \end{cases}$$

Corollary 5.2. *The coefficients $d_i(\delta)$ of (6) are nonnegative for all δ satisfying $0 \leq \delta \leq \gamma_q - \Delta'$. Moreover, if $\Delta' = \Delta^*$, then $M(y') = 0$ and $\theta(\alpha') > 0$.*

An immediate consequence of Theorem 5.1 and Corollary 5.2 is the fact that the set of all facets intersected by $y(\Delta)$ for $\Delta = \Delta^*$ can be identified by means of its correspondence with the set of all solutions to equation (6) in which

$$z_i = \begin{cases} 0 \text{ or } 1 & \text{if } d_i(0) = 0\,, \\ 0 & \text{otherwise,} \end{cases}$$

subject to the stipulation that $\Delta' = \Delta^*$.

Corollary 5.3. *Let α' be determined as in Theorem 5.1 for $\Delta' = \Delta^*$, and let δ_1 be the least positive value of δ such that $d_i(\delta) = \delta\, \theta(\alpha')$ for some $i \in P$. Then $\Delta = \Delta^* + \delta_1$ is the first intersection value of Δ beyond Δ^*,* *and the set of facet hyperplanes intersected for this value of Δ is given by the set of solutions to (6) which satisfy*

$$z_i = \begin{cases} 1 & \textit{for exactly one } i \textit{ such that } d_i(\delta_1) = \delta_1\,\theta(\alpha')\,, \\ 0 \textit{ or } 1 & \textit{if } d_i(\delta_1) = 0\,, \\ 0 & \textit{for all remaining } i\,. \end{cases}$$

The value δ_1 indicated in Corollary 5.3 can be instantly identified by reference to the definition of the coefficients $d_i(\delta)$ given in Theorem 5.1, showing that the second set of facet hyperplanes intersected by the edge $y(\Delta)$ can be determined with very little effort once Δ^* has been computed.

Corollary 5.4. *Let Δ', α' and δ_1 be given as in Corollary 5.3 and let δ_2 be the least positive value of $\delta > \delta_1$ such that either $d_i(\delta) = \delta\,\theta(\alpha')$ for some i, or $d_r(\delta) + d_s(\delta) = \delta\,\theta(\alpha')$ for some r and s. ($\delta_2 = \infty$ if neither of these conditions can be met.) If $\delta_2 \leq \gamma_u - \Delta'$, then $\Delta = \Delta' + \delta_2$ is the second intersection value of Δ beyond Δ^*, and the set of facet hyperplanes intersected for this value of Δ is given by the set of all solutions to (6) which satisfy the conditions of Corollary 5.3 with δ_2 in place of δ_1, together with those which satisfy $z_r = z_s = 1$ for some r and s such that $d_r(\delta_2) + d_s(\delta_2) = \delta_2\,\theta(\alpha')$ and for $i \neq r, s$,*

$$z_i = \begin{cases} 0 \textit{ or } 1 & \textit{if } d_i(\delta_2) = 0\,, \\ 0 & \textit{otherwise.} \end{cases}$$

If $\delta_2 > \gamma_q - \Delta'$, then redefine $\Delta' = \gamma_q$ (determining a new y' and α') and let δ_2 be the least positive value of δ such that $d_i(\delta) = M(y') + \delta\,\theta(\alpha')$ for some $i \in P$. Then $\Delta = \Delta' + \delta_2$ is the second intersection value of Δ beyond Δ^*, and the set of facet hyperplanes intersected for this value of Δ is given by the set of all solutions to (6) which satisfy

$$z_i = \begin{cases} 1 & \textit{for exactly one } i \textit{ such that } d_i(\delta_2) = M(y') + \delta_2\,\theta(\alpha'), \\ 0 \textit{ or } 1 & \textit{if } d_i(\delta_2) = 0, \\ 0 & \textit{for all remaining } i. \end{cases}$$

The value δ_2 of Corollary 5.4 requires only slightly more effort to determine than the value δ_1 of Corollary 5.3, and both of these values together should typically require less effort to determine than the value of Δ^*. The determination of δ_r and δ_s in Corollary 5.4 can be speeded by the observation that attention can be restricted to the subset of indexes i such that $d_i(\delta_1) \leq \delta_1\, \theta(\alpha')$, and by the observation that all i for which $d_i(\delta_1) < \delta_1\, \theta(\alpha')$ can be ignored except the one (or ones) with minimum nonzero value of $2h_i\, |b_{ij}|$.

Also, it is possible for the fourth and fifth intersection values to be determined relative to the same considerations that govern the determination of δ_2, provided the conditions of the last half of Corollary 5.4 continue to hold (implying that each subsequent calculation starts with a new Δ', ending finally with a calculation based on the conditions of the first half of Corollary 5.4).

References

[1] E. Balas, The intersection cut – A new cutting plane for integer programming, Management Science Research Report No. 187, Carnegie–Mellon Univ., Pittsburgh, Pa. (October 1969); Operations Res. 19 (January–February 1971) 19–39.

[2] E. Balas, Integer programming and convex analysis, Management Science Research Report No. 246, Carnegie–Mellon Univ., Pittsburgh, Pa. (April 1971).

[3] E. Balas, Ranking the facets of the octahedron, Management Science Research Report No. 252, Carnegie–Mellon Univ., Pittsburgh, Pa. (May 1971).

[4] E. Balas, V.J. Bowman, F. Glover and D. Sommer, An intersection cut from the dual of the unit hypercube, WP # 90-69-7, Carnegie–Mellon Univ., Pittsburgh, Pa. (April 1970); Operations Res. 19 (January–February 1971) 40–44.

[5] F. Glover, Convexity cuts and cut search, MSRS 70-2, School of Business, Univ. of Colorado, Boulder, Colo. (July 1970); Operations Res. 21 (January–February 1973).

[6] F. Glover, Cut search methods in integer programming, MSRS 70-3, School of Business, Univ. of Colorado, Boulder, Colo. (September 1970); Math. Programming 3 (1972) 86–100.